Research paper

Internet- and mobile-based intervention for depression in adults with chronic back pain: A health economic evaluation

Sarah Paganini a,*, Yannik Terhorst b, Lasse Bosse Sander c, Jiaxi Lin d, Sandra Schlicker e, David Daniel Ebert f, Matthias Berking e, Heleen Riper g, Harald Baumeister b

a Department of Sport Psychology, Institute of Sports and Sport Science, University of Freiburg, Germany
b Department of Clinical Psychology and Psychotherapy, Institute of Psychology and Education, Ulm University, Germany
c Department of Rehabilitation Psychology and Psychotherapy, Institute of Psychology, University of Freiburg, Germany
d Department of Psychiatry and Psychotherapy Medical Center, Faculty of Medicine, Freiburg, Germany
e Department of Clinical Psychology and Psychotherapy, Friedrich-Alexander University of Erlangen-Nuremberg, Germany
f Department of Psychology and Digital Mental Health Care, Technical University of Munich, Germany
g Department of Clinical, Neuro and Developmental Psychology, VU University Amsterdam, Netherlands

ARTICLE INFO

Keywords:
Depressive disorder
Chronic pain
Comorbidity
E-Health
Cost-effectiveness
Cost-utility

ABSTRACT

Background: Depression and comorbid chronic back pain (CBP) lead to high personal and economic burden. Internet- and mobile-based interventions (IMI) might be a cost-effective adjunct to established interventions.

Methods: A health economic evaluation was embedded into an observer-blinded, multicenter RCT (societal and health care perspective). We randomly assigned participants (≥18 years) with CBP and diagnosed depression from 82 orthopedic clinics across Germany to intervention (IG + treatment as usual [TAU]) or TAU control group (CG). The IG received a guided IMI. Primary outcomes were depression response and quality-adjusted life years (QALYs) at 6-months follow-up. Multiple imputation was used to address missing data. Incremental cost-effectiveness/cost-utility ratios (ICER/ICUR) and the probability of being cost-effective at different willingness-to-pay thresholds were calculated. Statistical uncertainty was estimated using bootstrapping techniques (N = 10,000).

Results: Between October 2015 and July 2017 210 participants were randomly assigned to IG (n = 105) and CG (n = 105). Depression response did not differ significantly between groups. QALYs were significantly higher in the IG compared to the CG. Taking the societal perspective and assuming a commonly used willingness-to-pay of €34,000/QALY, the intervention’s likelihood of being cost-effective was 64%.

Limitations: The main limitation is that the study was powered to detect clinical but not health economic differences between groups.

Conclusion: The IMI is considered cost-effective (vs. CG) for individuals with depression and CBP (societal perspective). These results are promising when considering the high individual and economic burden of this patient group. Further research is needed to adequately inform political decision makers before implementation into routine care.

1. Introduction

Major depression and chronic back pain (CBP) are two common conditions with a lifetime prevalence of up to 21% and 30%, respectively (Kessler and Bromet, 2013; Neuhauser et al., 2005). Both conditions are listed under the top five leading causes of global years lived with disability (Vos et al., 2017) and cause a high economic burden for society (Dagenais et al., 2008; Gaskin and Richard, 2012; Sobocki et al., 2006). Depression and CBP often occur comitantly (Bair et al., 2003; Harter et al., 2007). This further increases impairment in quality of life and healthcare costs (Baumeister et al., 2011, 2012).

It has been shown that psychological interventions, including cognitive-behavioral therapy (CBT) for CBP patients with depression or depressive symptoms, have a beneficial impact on depressive symptoms,
health-related quality of life and return to work, in addition to also improving pain intensity and pain-related interference (Hoffman et al., 2007; Williams et al., 2012). However, access to evidence-based psychotherapeutic interventions is still limited due to structural and attitudinal barriers such as long distances and waiting times or fear of stigmatisation (Andrade et al., 2014; Ebert et al., 2018; Ehde et al., 2014; Saxena et al., 2007). Internet- and mobile-based interventions (IMI) could contribute to overcome these barriers. With regard to CBP patients who often face problems in mobility, the delivery of CBT interventions via the internet could help to improve health for this patient group. Furthermore, IMI have been shown to be effective in the treatment of depression as well as chronic (low) back pain (Andrews et al., 2018; Macea et al., 2010). Cost-effectiveness is a further and often postulated advantage of IMI (Tate et al., 2009), as lower therapist costs might result in cost savings compared to face-to-face therapy. However, as CBP patients with depressive symptoms are seeking mental health care less likely than individuals with depressive symptoms alone (Bao et al., 2003; Carey et al., 2009), it is not clear what the economic impact of increased mental health care would be from a societal perspective or a health care perspective. There are still only few empirical studies of high methodological quality that contribute to clarify this point. Moreover, these studies draw contradicting conclusions about the cost-effectiveness of IMI in general and for depression in particular (Donker et al., 2015; Kolovos et al., 2018; Mitchell et al., 2021; Paganini et al., 2018). Therefore, the aim of this study is to examine the cost-effectiveness and cost-utility of a guided IMI for depression in CBP patients (eSano BackCare-D) against a treatment as usual (TAU) control group (CG). The IMI was provided as an orchestrated aftercare intervention. This health economic evaluation is conducted alongside a recently finalized, pragmatic, multicenter randomized controlled trial (RCT) that examined the effectiveness of eSano BackCare-D (Baumeister et al., 2021; Lin et al., 2017). The intervention group (IG) showed no significant reduction in depression severity at 6-month follow-up compared to the CG (Baumeister et al., 2021). However, there were significant group differences in the secondary outcomes depression remission and in health-related quality of life in favor of the IG at 6-month follow-up (Baumeister et al., 2021). Individuals in the CG might have compensated lack of adequate treatment by using more alternative therapies. Thus, in the present study it is hypothesized that the guided IMI is cost-effective compared to TAU.

2. Methods

2.1. Study design

This study is a health economic evaluation alongside a multicenter RCT. The analyses were conducted with a 6-month time horizon from a societal perspective and a health care perspective. Cost-effectiveness and cost-utility of a guided IMI for individuals with diagnosed depression and CBP was evaluated compared to TAU after orthopedic care. The health economic evaluation was conducted according to the International Society For Pharmacoeconomics and Outcomes Research (ISPOR) guidelines (Ramsey et al., 2015) and reported in accordance with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (Huser et al., 2013). The study has been registered at the WHO International Clinical Trials Registry (DRKS00009272). The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human subjects/patients were approved by the local ethics authorities (REC No 8022-6-BW-H-2015; No 297/14, 150513). Detailed information of the trial design and results concerning the clinical effectiveness can be found in the study protocol and the effectiveness paper (Baumeister et al., 2021; Lin et al., 2017).

2.2. Participants

Participants were recruited between October 2015 and July 2017 in 82 German orthopedic clinics and were randomly allocated (ratio of 1:1, variable block sizes of four, six and eight) to the IG or to the CG. An independent researcher performed the randomization using a web-based randomization program. Inclusion criteria were: (1) 18 years and older, (2) clinician-diagnosed CBP, (3) diagnosed mild to moderate depressive episode, dysthymia or persistent depressive disorder, (4) sufficient German language skills, and (5) internet access. Exclusion criteria were (1) receiving ongoing psychotherapy or psychotherapy within the past three months, (2) current suicidality or a suicidal attempt within the past five years, or (3) severe depression. Written informed consent was obtained from all subjects.

2.3. Intervention

The guided IMI eSano BackCare-D is a depression intervention based on CBT, adapted for individuals with CBP. It consists of six regular, three optional and two booster sessions. Participants were advised to work on one session per week (duration M = 54 min; SD = 23.7) and received semi-structured, written feedback from a psychologist (‘eCoach’). On average, the eCoach spent 101 min (SD = 38.4) per participant for guidance. There was no specific protocol for TAU, but health care utilization was assessed in detail with the Trimbos Institute and Institute of Medical Technology Questionnaire for Costs Associated with Psychiatric Illness (TiC-P) (Hakkaart-van Roijen et al., 2002). All participants had unrestricted access to TAU.

2.4. Outcome measures

Structured Clinical Interviews for DSM-IV/DSM-5 (SCID) (First et al., 1997, 2015) and assessment of depression severity were conducted by telephone at baseline (T0), nine weeks (T1) and 6-months follow-up (T2). Self-report measures were also collected at T0, T1 and T2 using a secured, internet-based assessment system (AES, 256-bit encrypted).

2.4.1. Cost-effectiveness analysis: primary outcome

The primary clinical outcome for the cost-effectiveness analysis was response in depression severity, defined by means of the reliable change index according to Jacobson and Truax (1991). Depression severity was assessed with the Hamilton Rating Scale for Depression (HAM-D; total scale score 0–52; higher scores indicate more (severe) depression scores) (Hamilton, 1960). The HAM-D contains 17 items and has good psychometric properties (Rehm and O’Hara, 1985). Participants were classified as responders if they showed a reliable change (greater than 6.88 points) (Jacobson and Truax, 1991) in the HAM-D at T2.

2.4.2. Cost-utility analysis: primary outcome

The primary outcome for the cost-utility analysis were quality-adjusted life-years (QALYs) based on the AQol-6D (total scale 20–99; lower scores indicate better health status) (Richardson et al., 2012) that has shown to be a reliable and valid instrument (Richardson et al., 2012, 2014). Participants completed the AQol-6D at T0, T1, and T2. This questionnaire comprises 20 items, which load on six dimensions (independent living, relationships, mental health, coping, pain and senses). Utility scores are anchored on 0 (meaning “a state as bad as being dead”) and 1 (meaning “full health”) (Drummond et al., 2015).

The EQ-5D-5L (dimensions: mobility, self-care, usual activities, pain/discomfort, anxiety/depression) is a valid and widely used instrument to calculate QALYs (Janssen et al., 2013) and was therefore used in the sensitivity analysis to assess robustness of the results. The German tariff was used to calculate index scores (van Reenen and Janssen, 2015).
2.5. Resource use and costing

Self-reported healthcare use and information about productivity loss in the past three months were assessed at T0 and T2 with the TiC-P (Hakkaart-van Roijen et al., 2002). Costs were calculated in Euros (€) for the reference year 2016 based on the German consumer price index (Statistisches Bundesamt, 2018). For unit costs see Supplement 1. There was no to discount costs or outcomes since the follow-up period was six months (Graf von der Schulenburg et al., 2008). To calculate the 6-month accumulated per-participants costs, the area under curve (AUC) method was used by linearly interpolating 3-month costs that were measured at T0 and T2 (Matthews et al., 1990).

\[ \text{AUC} = \frac{\text{Comp. T0} + \text{Comp. T2}}{2} \times 3 + \text{Costs T2} \]

2.5.1. Healthcare costs

Costs for inpatient care (e.g. hospital stays), outpatient care (physicians and psychological services) and costs for therapeutic interventions and appliances (e.g. physiotherapy and walking aid, respectively) were calculated according to German guidelines for unit cost prices from a societal perspective (Bock et al., 2015b).

2.5.2. Medication

The costs of medication were derived from the Lauer-Taxe (Lauer-Fischer GmbH, n.d.), a German encyclopedia for pharmaceutical professional groups and medical and health insurances. Exact prices were obtained in case the participant had provided the central pharmaceutical number (CPN). Pharmacy’s and manufacture’s discounts were taken into account (Bock et al., 2015a). If the CPN was not specified, the mean costs of the three largest packages with the same agent were calculated, discounts were subtracted and the result was weighted by the statutory population share (Bock et al., 2015a; Krauth, 2010).

2.5.3. Intervention costs

Intervention costs of €299 were based on the current market price for similar IMI (guided, with the same number of regular sessions; https://hellobetter.de/). This tariff includes all costs for developing and hosting the intervention.

2.5.4. Patient and family costs

Out-of-pocket costs (e.g. over-the-counter drugs) and direct non-medical costs (e.g. travel expenses caused by health-care service visits) were directly obtained from participants. For traveling by car each kilometer was valued at €0.30 (Bundesministerium der Justiz, 2008). If participants used public transport or taxi, they were asked to report the ticket price. In case of missing information, a weighted average price per medical appointment was assumed. Opportunity costs (e.g. time spent at the practitioners waiting room) were estimated at €22.67 per hour (Bock et al., 2015b). For paid domestic help, costs of €19.83 per hour were estimated (Bock et al., 2015b). Costs for unpaid domestic help were valued as opportunity costs as it was assumed that those who help offered their leisure time (Bock et al., 2015a).

2.5.5. Costs of productivity losses

Absenteism costs due to absence from work were calculated according to the human capital approach (Drummond et al., 2015). The participants’ self-reported number of lost work days was multiplied by the gross average of their income per day. The calculation of presenteeism costs was based on the Osterhaus method (Osterhaus et al., 1992). Participants reported the number of days with reduced functioning at work, which were weighted by an inefficiency score.

2.6. Statistical analysis

All analyses were based on the intention-to-treat principle. The study was powered to detect a mean group difference of \( d = 0.39 \) in the primary outcome HAM-D at \( p < 0.05 \) (2-tailed t-test) with a power of \( (1 - \beta) = 0.80 \) at T1. The present paper is a secondary analysis focusing on the cost-effectiveness and cost-utility. For the analyses missing cost and outcome data were imputed on the item-level for each timepoint by using multiple imputation by chained equations (MICE) (R package mice, version 3.3.0) (van Buuren and Groothuis-Oudshoorn, 2011). Average percentage of missingness across all variables was \( M = 14.5\% \) (SD = 12.23%, Min = 0%, Max = 57%). A missing at random (MAR) mechanism was assumed (Enders, 2010) and predictive mean matching was used as imputation method (Little, 1988; van Buuren and Groothuis-Oudshoorn, 2011). Following the strategy of van Buuren et al. (1999) the predictor matrix for imputation contained 1) all variables included in the analyses, 2) variables related to non-response (i.e., based on pairwise correlation between variables and response indicators) and 3) variables explaining variance (i.e., based on pairwise correlations) (van Buuren and Groothuis-Oudshoorn, 2011). The number of imputed data sets was set to \( m = 20 \) (Graham et al., 2007). Data aggregation (e.g., calculation of aggregated cost and utility variables) were performed within each of the 20 imputed datasets. Afterwards analysis was conducted separately on each imputed dataset and results were pooled using Rubin’s rule (Rubin, 1987; Rubin, 1996). Degrees of freedom were adjusted according to Barnard and Rubin (1999). The software R was used for all analyses (R Core Team, 2017). All analyses of clinical outcomes were conducted and reported in accordance with the CONSORT statement (Moher et al., 2010; Zwarnenstein et al., 2008) and can be found elsewhere (Baumeister et al., 2021).

The study was not powered to statistically test differences in health economic outcomes. Therefore, a probabilistic decision-making approach was adopted that accounts for the uncertainty of all measured parameters (Briggs and Gray, 1999; van Hout et al., 1994). Incremental cost-effectiveness ratios (ICER) and cost-utility ratios (ICUR) were calculated by dividing incremental costs (i.e. differences in costs over the 6-months period between individuals in the IG and the CG) by incremental effects (differences in effects; i.e. response or QALYs, respectively): ICER/ICUR = (Cost\text{IG} - Cost\text{CG}) / (Effect\text{IG} - Effect\text{CG}). To estimate QALY health gains during the 6-months follow-up period, the AUC was calculated using linear interpolation of utility scores between measurement points (Matthews et al., 1990). Group differences for response and QALYs were evaluated by using seemingly unrelated regression models. Baseline adjustments were made for the AQoL-6D and EQ-5D-5L.

To account for stochastic uncertainty in the ICER/ICUR, non-parametric bootstrapping by resampling patient-level data with 10,000 replications was applied. The number of bootstrap replications was based on the convergence of bootstrap estimates for bias and standard errors according to Briggs and colleagues (Briggs et al., 1997). Seemingly unrelated regression equation models (R package systemfit, version 1.1-22) (Henningsen and Hamann, 2007) were used to allow for correlated residuals of the cost and effect equations. In the presence of skewed variables and ratios, the non-parametric percentile method was used to obtain confidence intervals (van Hout et al., 1994). The generated ICER/ICUR were illustrated on cost-effectiveness planes. Cost-effectiveness acceptability curves were graphed to demonstrate the probability of the IMI being cost-effective compared to the CG depending on different WTP ceilings. To enhance comparability with other studies, the probabilities of commonly used WTP thresholds per QALY gained of the UK (£20,000 to £30,000; in Euro:~€23,000 to €34,000; McCabe et al., 2008) and the US (US$50,000; in Euro: ~€44,000; converted according to the European Central Bank; European Central Bank, n.d.) were calculated.
**Table 1**

Multiply imputed pooled effects and group differences after six months.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>IG (n = 104) Mean (SD)</th>
<th>CG (n = 105) Mean (SD)</th>
<th>Adjusted between group difference (95% CI)*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response (HAM-D)</td>
<td>0.365 (0.48)</td>
<td>0.324 (0.47)</td>
<td>0.042 (0.089 to 0.172)</td>
<td>0.529</td>
</tr>
<tr>
<td>QALY (AQoL-6D)</td>
<td>0.293 (0.09)</td>
<td>0.273 (0.09)</td>
<td>0.024 (0.007 to 0.041)</td>
<td>0.006†</td>
</tr>
<tr>
<td>Sensitivity analysis</td>
<td>0.361 (0.09)</td>
<td>0.356 (0.09)</td>
<td>0.024 (0.006 to 0.043)</td>
<td>0.01‡</td>
</tr>
</tbody>
</table>

IG, intervention group; CG, control group; SD, standard deviation; QALY, quality-adjusted life-year.

* Standardized regression estimate (odds ratio) adjusted for trial center, age, sex and baseline. Confidence intervals are based on robust standard errors.
† Statistically significant difference at p < 0.050.

2.7. Sensitivity analysis

A sensitivity analysis was conducted to assess the robustness of the outcomes of the main analysis. The EQ-5D-5L (Janssen et al., 2013) was used to calculate QALYs in addition to the AQoL-6D. The EQ-5D-5L constitutes a valid extension to the widely used EQ-5D-3L (EuroQol Group, 1990) with reduced ceiling effects and improved discriminatory power (Janssen et al., 2013).

3. Results

3.1. Sample characteristics

210 participants were randomized to either the IG (n = 105) or the CG (n = 105). One person of the IG withdrew the informed consent after randomization and was therefore excluded from all analyses. The average age was 50 years (SD = 9.36) and 60% of the participants were women (n = 125). 86% of the participants had a low (n = 132) or medium (n = 47) education level (Schneider, 2013). Participants’ characteristics and the flow chart are presented in detail elsewhere (Baumeister et al., 2021).

Dropout at T2 was 14.3% for the telephone assessment (IG: 16/104; CG: 14/105) and 25.8% for the online assessments (IG: 28/104; CG: 26/105). Dropout rates did not differ significantly between groups (telephone assessment: χ2(1) = 0.051, p = 0.822; online assessment: χ2(1) = 0.040, p = 0.842). Of the 104 individuals in the IG, 60.6% completed at least five sessions of the IMI (out of six regular and three optional sessions) and were defined as adherent.

3.2. Clinical outcomes

Primary outcomes and group differences at T2 are presented in Table 1. 38 individuals (36.5%) in the IG and 34 in the CG (32.4%) showed treatment response. The multiply imputed and pooled outcomes for response did not differ statistically between groups (β = 0.042, 95% CI −0.089 to 0.172, p = 0.529). Between-group differences in both AQoL-6D and EQ-5D-5L QALY gains were statistically significant in favor of the IG (AQoL-6D: β = 0.024, 95% CI −0.007 to 0.041, p = 0.006; EQ-5D-5L: β = 0.024, 95% CI −0.006 to 0.043, p = 0.01).

3.3. Cost outcomes

At baseline, mean total costs were €10,806 (SD = 6929) in the IG and €10,705 (SD = 6377) in the CG. Mean healthcare costs at T0 were €6887 (SD = 9485) in the IG and €5596 (SD = 6922) in the CG. Six-months accumulated and pooled per-participant costs for all cost categories as well as cost differences are presented in Table 2. In-patient care, rehabilitation, opportunity costs and with the highest values absenteeism costs were the greatest contributors to overall costs in both groups. Mean total societal costs were only slightly higher in the IG (€14,001) compared to the CG (€13,931). Opportunity costs (i.e. for waiting time before treatment) and absenteeism costs were considerably lower in the IG.

**Table 2**

Accumulated and pooled per-participants costs in €.

<table>
<thead>
<tr>
<th>Cost categories</th>
<th>IG (n = 104) Mean, € (SD)</th>
<th>CG (n = 105) Mean, € (SD)</th>
<th>Incremental costs, € (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health care costs</td>
<td>Intervention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical practitioner</td>
<td>477 (475)</td>
<td>412 (470)</td>
<td>64 (−74 to 203)</td>
<td>0.357</td>
</tr>
<tr>
<td>Mental health care</td>
<td>297 (575)</td>
<td>257 (577)</td>
<td>39 (−150 to 229)</td>
<td>0.684</td>
</tr>
<tr>
<td>Other medical specialist</td>
<td>655 (767)</td>
<td>651 (587)</td>
<td>93 (−115 to 301)</td>
<td>0.376</td>
</tr>
<tr>
<td>In-patient care (hospital)</td>
<td>2260 (7336)</td>
<td>1567 (4465)</td>
<td>639 (−1206 to 2593)</td>
<td>0.471</td>
</tr>
<tr>
<td>Day care</td>
<td>805 (3227)</td>
<td>434 (2785)</td>
<td>371 (−549 to 1292)</td>
<td>0.426</td>
</tr>
<tr>
<td>Rehabilitation</td>
<td>2117 (1186)</td>
<td>2083 (1466)</td>
<td>34 (−363 to 431)</td>
<td>0.866</td>
</tr>
<tr>
<td>Therapeutic appliances</td>
<td>46 (111)</td>
<td>44 (119)</td>
<td>2 (−32 to 35)</td>
<td>0.921</td>
</tr>
<tr>
<td>Prescription medication</td>
<td>232 (379)</td>
<td>238 (349)</td>
<td>−6 (−124 to 112)</td>
<td>0.918</td>
</tr>
<tr>
<td>Total healthcare costs</td>
<td>266 (631)</td>
<td>277 (453)</td>
<td>−11 (−176 to 145)</td>
<td>0.889</td>
</tr>
<tr>
<td>Over-the-counter drugs</td>
<td>40 (75)</td>
<td>50 (100)</td>
<td>−10 (−41 to 20)</td>
<td>0.502</td>
</tr>
<tr>
<td>Domestic help</td>
<td>1042 (2021)</td>
<td>1207 (2673)</td>
<td>−165 (−832 to 503)</td>
<td>0.627</td>
</tr>
<tr>
<td>Opportunity costs</td>
<td>1812 (3649)</td>
<td>2262 (4727)</td>
<td>−450 (−1724 to 825)</td>
<td>0.487</td>
</tr>
<tr>
<td>Total healthcare costs</td>
<td>2520 (3931)</td>
<td>3247 (5561)</td>
<td>−727 (−2149 to 695)</td>
<td>0.314</td>
</tr>
<tr>
<td>Presenteeism</td>
<td>1303 (2316)</td>
<td>1291 (2530)</td>
<td>12 (−679 to 703)</td>
<td>0.972</td>
</tr>
<tr>
<td>Total societal costs</td>
<td>7018 (9478)</td>
<td>5596 (6922)</td>
<td>1422 (−1076 to 3919)</td>
<td>0.262</td>
</tr>
</tbody>
</table>

IG, intervention group; CG, control group; SD, standard deviation; CI, confidence interval.

* Statistically significant difference at p < 0.050.
© Contacts with the general practitioner, specialist in internal medicine or other.
† I.e. physiotherapist, occupational therapist.
‡ I.e. bandages or walking aids.
§ I.e. for waiting time before treatment.
Table 3

Results of the multiply imputed and bootstrapped cost-effectiveness and cost-utility analyses and the sensitivity analysis.

<table>
<thead>
<tr>
<th>Analysis and outcome</th>
<th>Incremental costs, € (95% CI)</th>
<th>Incremental effects (95% CI)</th>
<th>Mean ICER/ICUR (95% CI)</th>
<th>Distribution over the cost-effectiveness plane (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost-effectiveness</strong></td>
<td></td>
<td></td>
<td></td>
<td>NE      NW  SE  SW</td>
</tr>
<tr>
<td>Societal: Response (HAM-D)</td>
<td>101 (–3749 to 3914)</td>
<td>0.041 (–0.090 to 0.169)</td>
<td>2470 (–302,338 to 290,227)</td>
<td>37 15 36 12</td>
</tr>
<tr>
<td>Health care: Response (HAM-D)</td>
<td>1436 (–978 to 3942)</td>
<td>0.041 (–0.090 to 0.169)</td>
<td>35,032 (–281,160 to 303,447)</td>
<td>65 23 8 4</td>
</tr>
<tr>
<td><strong>Cost-utility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societal: AQoL-6D QALYs</td>
<td>101 (–3749 to 3914)</td>
<td>0.024 (0.007 to 0.041)</td>
<td>4205 (–175,399 to 291,571)</td>
<td>52 0 48 0</td>
</tr>
<tr>
<td>Health care: AQoL-6D QALYs</td>
<td>1436 (–978 to 3942)</td>
<td>0.024 (0.007 to 0.041)</td>
<td>59,627 (–41,496 to 307,291)</td>
<td>88 0 12 0</td>
</tr>
<tr>
<td><strong>Sensitivity analysis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societal: EQ5D-5L</td>
<td>101 (–3749 to 3914)</td>
<td>0.024 (0.007 to 0.042)</td>
<td>4175 (–196,807 to 280,634)</td>
<td>52 0 48 0</td>
</tr>
</tbody>
</table>

CI, confidence interval; ICER, incremental cost-effectiveness ratio; ICUR, incremental cost-utility ratio; NE, north-east quadrant; NW, north-west quadrant; SE, south-east quadrant; SW, south-west quadrant; HAM-D, Hamilton Rating Scale for Depression; QALY, quality-adjusted life year.

Fig. 1. Cost-effectiveness plane and acceptability curve from a societal perspective (response).
Legend: a) Cost-effectiveness plane with the scatterplot of 10,000 replicates of the incremental cost-effectiveness ratio using mean differences in costs from a societal perspective and mean incremental effects (response) of eSano BackCare-D vs. treatment as usual. b) Cost-effectiveness acceptability curve showing the probability of eSano BackCare-D being cost-effective compared to treatment as usual at varying willingness-to-pay ceilings (based on 10,000 replicates of the incremental cost-effectiveness ratio using mean differences in costs from a societal perspective and mean incremental effects (response)).

Fig. 2. Cost-effectiveness plane and acceptability curve from a societal perspective (QALYs, based on AQoL-6D).
Legend: a) Cost-effectiveness plane with the scatterplot of 10,000 replicates of the incremental cost-effectiveness ratio using mean differences in costs from a societal perspective and mean incremental effects (QALYs, based on AQoL-6D) of eSano BackCare-D vs. treatment as usual. b) Cost-effectiveness acceptability curve showing the probability of eSano BackCare-D being cost-effective compared to treatment as usual at varying willingness-to-pay ceilings (based on 10,000 replicates of the incremental cost-effectiveness ratio using mean differences in costs from a societal perspective and mean incremental effects (QALYs based on AQoL-6D)).
IG. However, there were no statistically significant differences between the groups. The mean health care costs were higher in the IG as compared with the CG (Δ€1422). This can largely be explained by differences in in-patient care (Δ€639) and intervention costs (Δ€299).

3.4. Societal perspective

The results of the cost-effectiveness and cost-utility analysis are presented in Table 3. Cost-effectiveness analysis: On average the intervention led to similar health benefits in terms of individuals achieving response (incremental effect: 4.1%) and higher costs (incremental costs: €101) compared to the CG. Fig. 1a shows the cost-effectiveness plane representing 10,000 bootstrap replications. The majority of the bootstrapped ICER fell in the two eastern quadrants with the north-east quadrant (37%), indicating greater health gains at higher costs and the south-east quadrant (36%), indicating greater health gains at lower costs. At a WTP of €0 per response, the intervention’s probability of being cost-effective compared to TAU was 48%. For a probability of 50%, a WTP of €2280 per response would be necessary. At a WTP of about €55,000 the probability rises to 70% and stagnates around this level, independent of the WTP (Fig. 1b).

In the cost-utility analysis the IG reached higher health gains (incremental QALYs: 0.024) at higher costs (incremental costs: €101) compared to the CG (Table 3).

As shown in Fig. 2a, all of the bootstrapped ICUR fell in the eastern quadrants indicating probabilities of 52% (north-east) and 48% (south-east) that the intervention produces greater health gains at higher costs compared to the CG or at lower costs, respectively. At a WTP of €0, the probability that the intervention can be regarded as more cost-effective than the CG is 48%. This probability rises to 50%, 80% or 90% when society is willing-to-pay €3730, €83,500 and €138,900, respectively (Fig. 2b).

Probabilities for commonly used WTP thresholds per QALY gained were 59% (WTP €23,000), 64% (WTP €34,000) and 68% (WTP €44,000).

3.5. Health care perspective

Cost-effectiveness analysis: From a health care perspective, the health benefit (incremental effect for treatment response: 4.1%) was also achieved at higher costs (€1436). The majority of the simulated ICER fell in the north-east quadrant (65%), indicating greater health gains at higher costs, followed by the north-west quadrant (23%), indicating lower health gains at higher costs. At a WTP of €0 per response, the intervention’s probability of being cost-effective compared to TAU was 12%. For a probability of 50%, a WTP of €34,300 per response would be necessary. The probability rises to 70% (at a WTP of about €20,000) and stagnates around this level. In the cost-utility analysis, 88% of simulated ICUR fell in the north-east quadrant with the remaining 12% in the south-east quadrant. At a WTP of €0, the probability of the intervention being cost-effective compared to the CG is 12%. This probability rises to 50%, 80% or 90% when society is willing-to-pay €59,250, €119,500 and €165,800, respectively. Probabilities for commonly used WTP thresholds per QALY gained were 24% (WTP €23,000), 32% (WTP €34,000) and 39% (WTP €44,000).

3.6. Sensitivity analysis

Using the EQ-5D-5L resulted in similar incremental health gains (incremental QALYs: 0.024) compared to the results using the AQoL-6D (Table 3).

The cost-effectiveness plane is demonstrated in Supplement 2. At a WTP of €0 the probability of the intervention being cost-effective compared to the CG is 48% from a societal perspective. To get a probability of 50%, 80% or 90%, society’s WTP would have to be €3730, €80,300 and €131,800 for a QALY gained, respectively (see Supplement 3). Probabilities for commonly used WTP thresholds per QALY gained were 59% (WTP €23,000), 64% (WTP €34,000) and 68% (WTP €44,000).

4. Discussion

This study evaluated the cost-effectiveness and cost-utility of a guided IMI for individuals with depression and chronic back pain in orthopedic aftercare compared to a TAU control group. Response rates, in terms of reliable change in depression severity, did not differ significantly between groups. Quality-adjusted life years were significantly higher in the IG compared to the CG. Hereby, between group difference in response rate (4.2%) and QALYs (0.024) were slightly higher as compared to results of an individual-participant data meta-analysis on the cost-effectiveness of IMI for depression (between group difference in response rate: 2%; in QALYs: 0.01) (Kolovos et al., 2018). After six months, total societal costs were higher in the IG compared to the CG (Δ€101) leading to an ICER of 2470 and an ICUR of 4205. Both cost-effectiveness and cost-utility analysis from a societal perspective revealed a probability of eSano BackCare-D for being cost-effective compared to the CG of 48% at WTP of €0 (per response and QALY gained, respectively). The sensitivity analysis using the EQ-5D-5L yielded similar incremental QALY gains compared to using the AQoL-6D. The distribution of bootstrapped ICUR over the cost-effectiveness plane was uniform, showing the robustness of the data.

For a probability of being cost-effective of 50%, society’s WTP would have to rise to €2280 per depression response and to €3730 per QALY gained. Compared to other IMI for depression these costs are within the range in which IMI were classified as being cost-effective. In recent systematic reviews, WTP ranged between €0 and €5000 per relevant change in depression severity at a probability of 50% of being cost-effective compared to different control groups (Paganini et al., 2018). In terms of QALYs, IMI for depression with a probability of being cost-effective between 2% to 71% revealed cost per QALY estimates of around €24,000 to €60,000 per QALY (Mitchell et al., 2021). Further systematic reviews indicated probabilities of IMI being cost-effective above 50% with €3088 to €37,352 per QALY gained (Donker et al., 2015; Paganini et al., 2018).

When taking commonly used WTP thresholds per QALY gained of the UK (McCabe et al., 2008) (€20,000 to €30,000; in Euro: ~€23,000 to €34,000) or the US (US$50,000; in Euro: ~€44,000; converted according to the European Central Bank) (European Central Bank, n.d.) as reference, the probability of being cost-effective compared to the CG is 64% (WTP (€34,000) and 67% (WTP €44,000). Based on the reported results eSano BackCare-D can be considered cost-effective compared to the CG when judged from a societal perspective. However, from a health care perspective with costs per QALY gained of €59,627 the intervention would rather be classified as not cost-effective.

Previous studies found that interventions for the treatment of depression tend to lead to higher direct (health care) costs but to lower indirect costs compared to control groups (Barrett et al., 2005; Kolovos et al., 2018; Paganini et al., 2018). In the current study this was also true and explains why the classification of being cost-effective strongly depends on the perspective (societal vs. health care). However, other studies evaluating the cost-effectiveness of IMI for (subthreshold) depression revealed only small differences in incremental societal and health care costs and thus almost similar ICER and ICUR, irrespective of the perspective (Kolovos et al., 2018). A possible explanation might be that individuals with CBP show high health care use independent of having comorbid depression or not (Baumeister et al., 2012; Degenais et al., 2009). This might also be reflected in the overall costs. They were about three to four times higher in this study as compared to costs in other health economic evaluations of depression IMI (Kolovos et al., 2018), possibly leading to a ceiling effect and reducing the chances of a relevant reduction in health care costs. An opposite effect could be, that due to the experience with a psychological intervention, individuals in the IG might have been encouraged to make more use of the health care
system, leading to higher health care costs.

The results of this study indicating cost-effectiveness from a societal perspective with similar or even more beneficial incremental health and cost outcomes when compared to IMI for individuals with depression alone (Kolovos et al., 2018) are remarkable due to several reasons. First, the co-occurrence of depression and chronic pain can lead to treatment resistance (Bair et al., 2003; Sullivan et al., 2008). Thus, it might be more difficult to improve depression and quality of life outcomes in individuals with depression and CBP. Second, there might be a higher challenge in reducing indirect costs in this specific population. It could be shown that indirect costs were not increased in individuals with CBP and mental disorders compared to individuals with CBP alone (Bauermeister et al., 2012). Furthermore, CBP is one of the main causes of work disability and long-term absence (Dagenais et al., 2008) and eSano BackCare-D did not primarily aim to reduce pain intensity. Third, treatment programs for chronic pain (such as the program conducted by all participants in this study) improve overall functioning (Turk, 2002) and have positive effects on return to work, medication use and general utilization of the health care system and thus can contribute to a reduction in costs (Turk, 2002). This could be a reason why previous studies showed that IMI for individuals who were already embedded into routine care have a lower probability of being cost-effective (Paganini et al., 2018).

Nevertheless, and despite these specificities of CBP as comorbidity and the setting of orthopedic aftercare, the intervention was able to contribute to (slightly) higher health gains and lower indirect costs in the intervention group.

4.1. Limitations

First, the health economic evaluation is a secondary analysis. Thus, this study was powered to detect clinical but not health economic differences between groups. This is a common problem in health economic evaluations alongside RCTs (Ramsey et al., 2015) and might reduce generalization. To handle this problem we relied on a probabilistic decision-making approach, thus accounting for uncertainty (Briggs and Gray, 1999; van Hout et al., 1994). Second, as clinical effectiveness seems to be an indicator of cost-effectiveness (Paganini et al., 2018), the exclusion of individuals with severe depression from this trial might have had a negative impact on the effectiveness of the IMI and thus, in return also on the cost-effectiveness results. Third, all analyses relied on a 6-months follow-up period. Thus, no conclusions about long-term cost-effectiveness can be drawn. However, as utility scores had converged only slightly over time and were higher within the IG at 6 months these cost-effectiveness estimates can be considered as rather conservative. Fourth, resource use and productivity losses were assessed via self-report. However, the questionnaire used in this study has been shown to be valid for recall periods up to three months, reducing recall bias (van den Brink et al., 2005).

5. Conclusions

This health economic evaluation of a guided IMI for individuals with depression and CBP revealed good value-for-money depending on the health economic perspective and on societies’ willingness-to-pay. When compared to commonly used willingness-to-pay thresholds (i.e. €23,000 or €44,000 per QALY gained) the probability of being cost-effective of 59% and 67% at these WTP indicate that the intervention can be classified as cost-effective from a societal perspective. Due to the high individual and economic burden caused by depression and CBP and by considering the specific challenges that this group brings with it, it is a promising result that health economic outcomes are similar or even more beneficial compared to other IMI for depression only (Kolovos et al., 2018). However, as this is the first study evaluating the cost-effectiveness of a depression IMI in this target group and in this setting, implementation into orthopedic aftercare should be preceded by further research.

CRediT authorship contribution statement

HB is the overall guarantor of this report. HB, DDE, MB and HR conceived and obtained funding for the RCT underlying this health economic evaluation. SP, LBS, JL, SaS, HB, and DDE contributed to the design of the study and SP, HB and DDE contributed to the design of this health economic evaluation. SP, LBS, JL, SaS and HB adapted the intervention content and the assessment. SP, LBS, JL and SaS were responsible for recruitment. SaS was responsible for randomization and allocation as well as study administration. SP and YT were responsible for data analysis. SP had full access to all data in the study and takes responsibility for the integrity and accuracy of the data analysis. SP wrote the draft of this manuscript. All authors revised the manuscript and approved its final version.

Declaration of competing interest

All authors of the manuscript were involved in the development of eSano BackCare-D or its predecessor versions. A committee of independent scientists has been formed (DSMB) to supervise study-related decisions and prevent any influence of potential conflicts of interest.

LBS and SaS have received payments for workshops on e-mental-health. HB received consultancy fees, reimbursement of congress attendance and travel costs as well as payments for lectures from Psychotherapy and Psychiatry Associations and as well as Psychotherapy Training Institutes in the context of E-Mental-Health topics. He has been the beneficiary of study support (third-party funding) from several public funding organizations. DDE and MB possess shares in the GET. On Institut GmbH, which works to transfer research findings on Internet- and mobile-phone-based health interventions into routine care. DDE has received payments from several companies and health insurance providers for advice on the use of internet-based interventions. He has received payments for lectures from Psychotherapy and Psychiatry Associations and has been the beneficiary of third-party funding from health insurance providers. DDE and MB were not involved in the data analysis.

Acknowledgements

We thank our many study assistants for their support in recruitment and data processing. We thank Maren Clarissa Köhler and Amelie Schlawinsky for their support in data analysis and literature research for the manuscript. Many thanks go to Claudia Buntrock and Fanny Kähkle for sharing their expertise in health economic evaluations. We thank Ellen Meierotto for supervising the SCID interviewers. Many thanks go to the Data Safety and Monitoring Board, consisting of Prof Dr. Martin Hautzinger, Prof Dr. Martin Härtler and Dr. Levente Kriston as well as the Clinical Trials Unit Freiburg for their support in carrying out the study project.

Special thanks go to the cooperating orthopedic rehabilitation clinics Schoen Klinik, Bad Staffelstein; Rehaklinik Sonnhalde, Donaueschingen; RehaKlinikum Bad Säckingen; Städt. Rehakliniken Bad Waldsee; Schwarzwaldklinik Orthopädie - Abteilung Medizinische Rehabilitation, Bad Krozingen; Rheintalklinik Bad Krozingen; REGIO-Reha Tagesklinik, Freiburg and Universitäts- und Rehabilitationskliniken, Ulm as well as all orthopedic rehabilitation clinics participating in the second recruitment wave.

Role of funding source

This study was funded by the German Federal Ministry of Education and Research (project “Effectiveness of a guided web-based intervention for depression in back pain rehabilitation aftercare”, grant number: 01GY1330A; 01GY1330B). The funders had no role in the study design.


